

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 08-146985

(43)Date of publication of application : 07.06.1996

(51)Int.Cl.

G10L 3/00

G10L 3/02

G11B 20/00

(21)Application number : 06-283641

(71)Applicant : SANYO ELECTRIC CO LTD

(22)Date of filing : 17.11.1994

(72)Inventor : TANAKA KOJI
MIYATAKE MASANORI
IIDA MASAYUKI

(54) SPEAKING SPEED CONTROL SYSTEM

(57)Abstract:

PURPOSE: To provide a speaking speed control system which can control the speaking speed on the basis of speaking speed control information for a reproducing device which previously stores information for controlling the speaking speed in sent data or on the sound recording media, etc., and receives and reproduces the transmitted data or a reproducing device for the sound recording media.

CONSTITUTION: The speaking speed control system is equipped with an editing device 1 which generates a signal by adding speaking speed control information to a speech signal and a reproducing device 2 which separates the signal generated by the editing device 1 into the speech signal and speaking speed control information and performs speaking speed control over the speech signal according to the speaking speed control information.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-146985

(43) 公開日 平成8年(1996)6月7日

(51) Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 L 3/00	H			
3/02	A			
G 1 1 B 20/00	Z	9294-5D		

審査請求 未請求 請求項の数9 O L (全 13 頁)

(21) 出願番号 特願平6-283641

(22) 出願日 平成6年(1994)11月17日

(71) 出願人 000001889

三洋電機株式会社

大阪府守口市京阪本通2丁目5番5号

(72) 発明者 田中 浩司

大阪府守口市京阪本通2丁目5番5号 三
洋電機株式会社内

(72) 発明者 宮武 正典

大阪府守口市京阪本通2丁目5番5号 三
洋電機株式会社内

(72) 発明者 飯田 正幸

大阪府守口市京阪本通2丁目5番5号 三
洋電機株式会社内

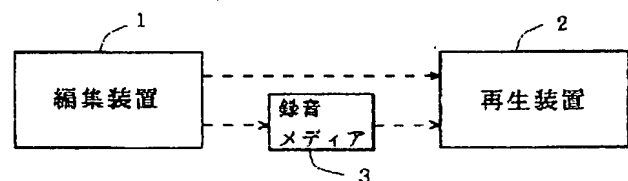
(74) 代理人 弁理士 香山 秀幸

(54) 【発明の名称】 話速制御システム

(57) 【要約】

【目的】 この発明は、予め送信データ、録音メディア等に話速を制御するための情報を入れておき、送信データを受信して再生する再生装置または録音メディアの再生装置において、話速制御情報に基づいて話速を制御できる話速制御システムを提供することを目的とする。

【構成】 話速制御システムにおいて、話速制御情報が音声信号に付加された信号を生成する編集装置1、および編集装置1によって生成された信号から音声信号と話速制御情報とを分離し、かつ音声信号を話速制御情報にしたがって話速制御する再生装置2を備えている。



【特許請求の範囲】

【請求項 1】 話速制御情報が音声信号に付加された信号を生成する編集装置、および編集装置によって生成された信号から音声信号と話速制御情報とを分離し、かつ音声信号を話速制御情報にしたがって話速制御する再生装置を備えている話速制御システム。

【請求項 2】 編集装置は、話速制御情報が音声信号に付加された信号を生成して、録音メディアに記録するものであり、再生装置は上記録音メディアから話速制御情報が音声信号に付加された信号を読み出して、音声信号と話速制御情報とを分離し、かつ音声信号を話速制御情報にしたがって話速制御するものである請求項 1 に記載の話速制御システム。

【請求項 3】 編集装置は、話速制御情報が音声信号に付加された信号を生成して、送信データを作成するものであり、再生装置は上記送信データを受信して、受信した上記送信データから音声信号と話速制御情報とを分離し、かつ音声信号を話速制御情報にしたがって話速制御するものである請求項 1 に記載の話速制御システム。

【請求項 4】 時分割多重または周波数分割多重方式によって、話速制御情報が音声信号に付加されることを特徴とする請求項 1、2 および 3 のいずれかに記載の話速制御システム。

【請求項 5】 切換伝送方式によって、話速制御情報が音声信号に付加されることを特徴とする請求項 1、2 および 3 のいずれかに記載の話速制御システム。

【請求項 6】 再生装置は、音声信号を時間軸圧縮伸長処理する処理手段と、音声信号を削除処理する処理手段とを有する話速変換部を備えており、話速制御情報は、音声信号を時間軸伸長圧縮処理を行なうための情報と、音声信号を削除処理するための情報とを少なくとも含んでいることを特徴とする請求項 1、2、3、4 および 5 のいずれかに記載の話速制御システム。

【請求項 7】 再生装置は、音声信号を時間軸圧縮伸長処理する処理手段と、音声信号を削除処理する処理手段とを有する話速変換部を備えており、

話速制御情報は、音声区間であることを示す情報と、無音区間の継続数を示す情報とを少なくとも含んでいることを特徴とする請求項 1、2、3、4 および 5 のいずれかに記載の話速制御システム。

【請求項 8】 音声信号を分析して話速制御情報を生成する手段、および生成された話速制御情報を音声信号に付加する手段を備えている編集装置。

【請求項 9】 話速制御情報が付加された音声信号から音声信号と話速制御情報とを分離する手段、および音声信号を話速制御情報にしたがって話速制御する手段を備えている再生装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 この発明は、放送、電話などの通信、パソコン、音声メール等の送信データに話速を制御するための制御情報を入れておき、受信側において制御情報に基づいて話速を制御する話速変換制御システム、または録音テープ、光磁気ディスク、レーザディスク、CD-ROM、ビデオCD、ICメモリ等の録音メディアに話速を制御するための制御情報を記録しておき、再生時に制御情報に基づいて話速を制御する話速制御システムに関する。

【0002】

【従来の技術】 本出願人は、音声区間の入力音声信号の時間長さを圧縮・伸長する手段および所定長さ以上の無音区間の入力音声信号を削除する削除手段を有する話速変換装置を開発している。このような話速変換装置では、入力された音声信号を分析して、分析結果に応じて圧縮・伸長処理、削除処理等が行なわれている。

【0003】 このような話速変換装置の入力音声信号としては、放送、電話などの通信、パソコン、音声メール等の送信データ、録音テープ、光磁気ディスク、レーザディスク、CD-ROM、ビデオCD、ICメモリ等の録音メディアから読み出されたデータがある。

【0004】

【発明が解決しようとする課題】 この発明は、予め送信データ、録音メディア等に話速を制御するための情報を入れておき、送信データを受信して再生する再生装置または録音メディアの再生装置において、話速制御情報に基づいて話速を制御できる話速制御システムを提供することを目的とする。

【0005】

【課題を解決するための手段】 この発明による話速制御システムは、話速制御情報が音声信号に付加された信号を生成する編集装置、および編集装置によって生成された信号から音声信号と話速制御情報とを分離し、かつ音声信号を話速制御情報にしたがって話速制御する再生装置を備えていることを特徴とする。

【0006】 編集装置としては、たとえば、話速制御情報が音声信号に付加された信号を生成して、録音メディアに記録するものが用いられ、再生装置としては、たとえば、上記録音メディアから話速制御情報が音声信号に付加された信号を読み出して、音声信号と話速制御情報とを分離し、かつ音声信号を話速制御情報にしたがって話速制御するものが用いられる。

【0007】 また、編集装置としては、たとえば、話速制御情報が音声信号に付加された信号を生成して、送信データを作成するものが用いられ、再生装置としては、たとえば、上記送信データを受信して、受信した上記送信データから音声信号と話速制御情報とを分離し、かつ音声信号を話速制御情報にしたがって話速制御するものが用いられる。

【0008】 話速制御情報を音声信号に付加する方法と

しては、時分割多重方式、周波数分割多重方式、切換伝送方式等が用いられる。

【0009】再生装置として、たとえば、音声信号を時間軸圧縮伸長処理する処理手段と、音声信号を削除処理する処理手段とを有する話速変換部を備えたものが用いられる。話速制御情報としては、たとえば、音声信号を時間軸伸長圧縮処理を行なうための情報と、音声信号を削除処理するための情報とが用いられる。また、話速制御情報としては、音声区間であることを示す情報と、無音区間の継続数を示す情報とが用いられる。

【0010】この発明による編集装置は、音声信号を分析して話速制御情報を生成する手段、および生成された話速制御情報を音声信号に付加する手段を備えていることを特徴とする。

【0011】この発明による再生装置は、話速制御情報が付加された音声信号から音声信号と話速制御情報とを分離する手段、および音声信号を話速制御情報にしたがって話速制御する手段を備えていることを特徴とする。

【0012】

【作用】この発明による話速制御システムでは、話速制御情報が音声信号に付加された信号が編集装置によって生成される。そして、再生装置側で、編集装置によって生成された信号から音声信号と話速制御情報とが分離され、かつ音声信号が話速制御情報にしたがって話速制御される。

【0013】この発明による編集装置では、音声信号を分析して話速制御情報が生成され、生成された話速制御情報が音声信号に付加される。

【0014】この発明による再生装置では、話速制御情報が付加された音声信号から音声信号と話速制御情報とが分離され、音声信号が話速制御情報にしたがって話速制御される。

【0015】

【実施例】以下、図面を参照して、この発明の実施例について説明する。

【0016】図1は、話速制御システムを示している。話速制御システムは、話速制御情報が音声信号に付加された信号を生成する編集装置1と、編集装置1によって生成された信号から音声信号と話速制御情報とを分離し、かつ音声信号を話速制御情報にしたがって話速制御する再生装置2とを備えている。

【0017】編集装置1としては、話速制御情報が音声信号に付加された信号から送信データを作成するもの、話速制御情報が音声信号に付加された信号を録音メディア3に記録するものがある。前者の場合には、編集装置1から話速制御情報が音声信号に付加された送信データが再生装置に送信される。後者の場合には、編集装置1によって録音メディア3から話速制御情報が音声信号に付加された信号が読み出される。

【0018】図2は、編集装置1の構成を示している。

編集装置1では、元となる音声信号が音声信号分析部11によって分析されて話速制御情報が生成される。生成された話速制御情報は、コード化部12によってコード化される。そして、元となる音声信号と生成されたコードとが多重化部13によって多重化される。この多重化信号に基づいて送信データが作成されるか、またはこの多重化信号が録音メディアに記録される。

【0019】図3は、再生装置2の構成を示している。再生装置2には、編集装置1によって作成された送信データまたは編集装置1によって多重化信号が記録された録音メディアからの読み出しデータが入力される。

【0020】入力された多重化信号は、復調部21で復調され、話速制御情報と音声信号とに分離される。音声信号は、話速制御部22に入力される。話速制御部22には、話速制御情報と復調部21からの音声／話速制御情報同期用制御信号が制御信号として入力される。話速制御部22では、話速制御情報に基づいて、入力音声の話速が制御される。

【0021】多重化方式としては、周波数分割多重方式、時分割多重方式等が用いられる。なお、編集装置1において、音声信号が符号化されている場合には、復調部21によって分離された音声信号は、復号化された後に話速制御部22に送られる。

【0022】図4は、話速制御部22の構成を示している。話速制御部22は、話速制御情報を解析する話速制御情報解析部31、音声／話速制御情報同期用制御信号に基づいて、話速制御情報解析部31に同期信号を送る制御情報同期化部32および話速制御情報解析部31の解析結果に基づいて音声信号を話速変換する話速変換部33を備えている。

【0023】話速変換部33は、話速制御情報解析部31の解析結果に応じて、入力音声信号の圧縮伸長処理、入力音声信号の削除処理等を行なう。

【0024】図5は、編集装置1の第1動作例を示している。ここでは、映像と音声とを伴う放送番組の音声信号が入力されているものとする。また、説明の便宜上、再生装置2側で設定される再生速度としては、1倍速再生と2倍速再生との2種があるとする。また、音声信号に多重化される話速制御情報には、1倍速再生用と2倍速再生用とがある。

【0025】まず、音声信号分析部11によって入力音声信号の所定区間毎に入力音声信号のパワー平均値Pが算出される(ステップ1)。次に、パワー平均値Pが所定のしきい値Th以上か否かが判別される(ステップ2)。パワー平均値Pが所定のしきい値Th以上($P \geq Th$)である場合には、当該区間は音声区間であると判別され、音声信号の時間長さを伸長する区間と判定される(ステップ3)。ここでは、音声区間以外の低レベルの定常雑音や環境音も無音区間として取り扱われるように、しきい値が設定されている。この後、圧縮率 α が設

定される（ステップ4）。

【0026】2倍速再生用の圧縮率 α は、たとえば、 $1/2 \leq \alpha \leq 1$ の範囲内の所定の値に設定される。圧縮率 α が2倍速再生時の一般的な圧縮率である $1/2$ である場合には、出力音声速度は入力音声速度の2倍となる。圧縮率 α が1である場合には、出力音声速度は入力音声速度の1倍となる。したがって、出力音声速度が入力音声速度の1倍以上で2倍以下となる範囲内で、圧縮率 α が設定される。

【0027】ここでは、圧縮率 α が $2/3$ に設定されるとする。この場合には、再生時において、音声の3ピッチ周期が2ピッチ周期に間引かれる。このため、再生時においては、出力音声速度は入力音声速度の $3/2$ 倍となる。このように、圧縮率 $2/3$ で圧縮された場合には、2倍速再生時の一般的な圧縮率 $1/2$ の場合に比べて、 $2/3 - 1/2 = 1/6$ だけ、音声が伸長されることになる。

【0028】1倍速再生用の圧縮率 α は、たとえば、 $1 \leq \alpha \leq 3/2$ の範囲内の所定の値に設定される。圧縮率 α が1倍速再生時において1である場合には、出力音声速度は入力音声速度の1倍となる。圧縮率 α が $3/2$ である場合には、出力音声速度は入力音声速度の $2/3$ 倍となる。したがって、出力音声速度が入力音声速度の $2/3$ 倍以上で1倍以下となる範囲内で、圧縮率 α が設定される。

【0029】ここでは、圧縮率 α が $3/2$ に設定されるとする。この場合には、再生時において、音声の2ピッチ周期が3ピッチ周期に伸長される。このため、再生時においては、出力音声速度は入力音声速度の $2/3$ 倍となる。このように、圧縮率 $3/2$ で伸長された場合には、1倍速再生の通常再生時に対して、 $3/2 - 1 = 1/2$ だけ、音声が伸長されることになる。

【0030】次に、1倍速再生速度および2倍速再生速度に応じてそれぞれ設定された圧縮率 α からなる話速制御情報が生成され、コード化部12によってコード化される（ステップ5）。次に、多重化部13によって、当該区間の入力音声信号と話速制御情報とが多重化される（ステップ6）。そして、この多重化信号に基づいて送信データが作成されるかまたはこの多重化信号が録音メディアに記録される（ステップ7）。

【0031】上記ステップ2において、パワー平均値Pが所定のしきい値 T_h より小さいときには（ $P < T_h$ ）、当該区間は無音区間であると判別され、無音区間の継続数が算出される（ステップ8）。そして、無音区間の継続数が所定数 T_{del} 以上であるか否かが判別される（ステップ9）。無音区間の継続数が所定数 T_{del} より少ないときには、音声信号の時間長さを伸長する区間と判定される（ステップ10）。

【0032】次に、上記ステップ4と同様に、1倍速再生用および2倍速再生用の圧縮率 α が設定される（ステ

ップ11）。そして、1倍速再生および2倍速再生用の圧縮率 α からなる話速制御情報が生成され、コード化部12によってコード化される（ステップ11）。次に、多重化部13によって、当該区間の入力音声信号と話速制御情報とが多重化される（ステップ6）。そして、この多重化信号に基づいて送信データが作成されるかまたはこの多重化信号が録音メディアに記録される（ステップ7）。

【0033】上記ステップ9において、無音区間の継続数が所定数 T_{del} 以上であると判定されたときには、当該区間は削除すべき区間と判定される（ステップ12）。そして、当該区間の入力音声信号を削除区間とする制御情報が生成され、コード化部12によってコード化される（ステップ13）。

【0034】次に、多重化部13によって、当該区間の入力音声信号と話速制御情報とが多重化される（ステップ6）。そして、この多重化信号に基づいて送信データが作成されるかまたはこの多重化信号が録音メディアに記録される（ステップ7）。

【0035】このようにして生成された多重化信号が再生装置2側で再生される場合には、操作者によって再生速度（1倍速または2倍速）が設定される。入力音声信号のうち、設定された再生速度に対する圧縮率 α が話速制御情報として設定されている区間では、その圧縮率 α で入力音声信号が時間軸圧縮伸長処理された後、再生される。また、入力音声信号のうち、話速制御情報によって削除される区間であると指定されている区間では、入力音声信号が削除される。

【0036】図6は、編集装置1の第2動作例を示している。ここでは、映像と音声とを伴う放送番組の音声信号が入力されているものとする。また、説明の便宜上、再生装置2側で設定される再生速度としては、1倍速再生と2倍速再生との2種があるとする。また、音声信号に多重化される話速制御情報には、1倍速再生用と2倍速再生用とがある。

【0037】まず、音声信号分析部11によって入力音声信号の所定区間毎に入力音声信号のパワー平均値Pが算出される（ステップ21）。次に、パワー平均値Pが所定のしきい値 T_h 以上か否かが判別される（ステップ22）。パワー平均値Pが所定のしきい値 T_h 以上（ $P \geq T_h$ ）である場合には、当該区間は音声区間であると判別され、音声信号の時間長さを伸長する区間と判定される（ステップ23）。そして、音声信号の時間長さを伸長する区間とする旨の話速制御情報が生成され、コード化部12によってコード化される（ステップ24）。

【0038】次に、多重化部13によって、当該区間の入力音声信号と話速制御情報とが多重化される（ステップ25）。この多重化信号に基づいて送信データが作成されるかまたはこの多重化信号が録音メディアに記録される（ステップ26）。

【0039】上記ステップ22において、パワー平均値 P が所定のしきい値 T_h より小さいときには ($P < T_h$)、当該区間は無音区間であると判別され、無音区間の継続数が算出される (ステップ27)。そして、無音区間の継続数を示す話速制御情報がコード化部12によって生成される (ステップ28)。

【0040】次に、多重化部13によって、当該区間の入力音声信号と話速制御情報が多重化される (ステップ25)。この多重化信号に基づいて送信データが作成されるかまたはこの多重化信号が録音メディアに記録される (ステップ26)。

【0041】このようにして生成された多重化信号が再生装置2側で再生される場合には、操作者によって再生速度 (1倍速または2倍速) が設定される。また、設定した再生速度に応じた圧縮率 α が操作者によって設定される。また、削除区間を決定するための継続長 T_{del} が操作者によって設定される。

【0042】そして、入力音声信号のうち、話速制御情報によって伸長する区間であると指定された区間では、操作者によって設定された圧縮率 α で入力音声信号が時間軸圧縮伸長処理される。また、入力音声信号のうち、話速制御情報によって無音区間の継続数が指定されかつその継続長が操作者によって決定された継続長 T_{del} 以上である区間では、入力音声信号が削除される。また、入力音声信号のうち、話速制御情報によって無音区間の継続数が指定されかつその継続長が操作者によって決定された継続長 T_{del} 未満である区間では、操作者によって設定された圧縮率 α で入力音声信号が時間軸圧縮伸長処理される。

【0043】図7は、編集装置1の第3動作例を示している。ここでは、映像と音声とを伴う放送番組の音声信号が入力されているものとする。また、説明の便宜上、再生装置2側で設定される再生速度としては、1倍速再生と2倍速再生との2種があるとする。また、音声信号に多重化される話速制御情報には、1倍速再生用と2倍速再生用とがある。

【0044】まず、音声信号分析部11によって入力音声信号の所定区間毎に入力音声信号のパワー平均値 P が算出される (ステップ31)。次に、パワー平均値 P が所定のしきい値 T_h 以上か否かが判別される (ステップ32)。パワー平均値 P が所定のしきい値 T_h 以上 ($P \geq T_h$) である場合には、当該区間は音声区間であると判別され、音声信号の時間長さを伸長する区間と判定される (ステップ33)。

【0045】次に、現時点での音声の伸長量、すなわち、再生時の入力信号に対する出力信号の遅延時間に応じた値が、1倍速再生速度および2倍速再生速度ごとに算出される (ステップ34)。そして、1倍速再生速度および2倍速再生速度ごとに、伸長量に応じた圧縮率 α が設定される (ステップ35)。

【0046】2倍速再生用の圧縮率 α は、たとえば、 $1/2 \leq \alpha \leq 1$ の範囲内で設定され、現時点での伸長量が小さいほど圧縮率 α は大きくされる。圧縮率 α が2倍速再生時の一般的な圧縮率である $1/2$ である場合には、出力音声速度は入力音声速度の2倍となり、圧縮率 α が1である場合には、出力音声速度は入力音声速度の1倍となる。したがって、出力音声速度が入力音声速度の1倍以上で2倍以下となる範囲内で、圧縮率 α が設定される。

【0047】また、現時点での伸長量が大いほど、出力音声速度が速くなるように圧縮率 α が設定される。この理由は、入力信号に対する出力信号の遅延時間が所定時間以上になるのを防止するためである。

【0048】1倍速再生用の圧縮率 α は、たとえば、 $1 \leq \alpha \leq 3/2$ の範囲内の所定の値に設定され、現時点での伸長量が小さいほど圧縮率 α は大きくされる。圧縮率 α が1倍速再生時において1である場合には、出力音声速度は入力音声速度の1倍となる。圧縮率 α が $3/2$ である場合には、出力音声速度は入力音声速度の $2/3$ 倍となる。したがって、出力音声速度が入力音声速度の $2/3$ 倍以上で1倍以下となる範囲内で、圧縮率 α が設定される。

【0049】1倍速再生用の圧縮率 α についても、現時点での伸長量が大いほど、出力音声速度が速くなるように圧縮率 α が設定される。この理由は、入力信号に対する出力信号の遅延時間が所定時間以上になるのを防止するためである。

【0050】次に、1倍速再生速度および2倍速再生速度に応じてそれぞれ設定された圧縮率 α からなる話速制御情報が生成され、コード化部12によってコード化される (ステップ36)。次に、多重化部13によって、当該区間の入力音声信号と話速制御情報が多重化される (ステップ37)。そして、この多重化信号に基づいて送信データが作成されるかまたはこの多重化信号が録音メディアに記録される (ステップ38)。

【0051】上記ステップ32において、パワー平均値 P が所定のしきい値 T_h より小さいときには ($P < T_h$)、当該区間は無音区間であると判別され、無音区間の継続数が算出される (ステップ39)。そして、無音区間の継続数が所定数 T_{del} 以上であるか否かが判別される (ステップ40)。無音区間の継続数が所定数 T_{del} より少ないときには、音声信号の時間長さを伸長する区間と判定される (ステップ41)。

【0052】次に、上記ステップ34と同様に、現時点での音声の伸長量、すなわち、再生時の入力信号に対する出力信号の遅延時間に応じた値が、1倍速再生速度および2倍速再生速度ごとに算出される (ステップ42)。そして、上記ステップ35と同様に、1倍速再生速度および2倍速再生速度ごとに、伸長量に応じた圧縮率 α が設定される (ステップ43)。

【0053】そして、1倍速再生速度および2倍速再生速度に応じてそれぞれ設定された圧縮率 α からなる話速制御情報が生成され、コード化部12によってコード化される(ステップ44)。次に、多重化部13によって、当該区間の入力音声信号と話速制御情報とが多重化される(ステップ37)。そして、この多重化信号に基づいて送信データが作成されるかまたはこの多重化信号が録音メディアに記録される(ステップ38)。

【0054】上記ステップ40において、無音区間の継続数が所定数 T_{del} 以上であると判定されたときには、当該区間は削除すべき区間と判定される(ステップ45)。そして、当該区間の入力音声信号を削除区間とする制御情報がコード化部12によって生成される(ステップ46)。

【0055】次に、多重化部13によって、当該区間の入力音声信号と話速制御情報とが多重化される(ステップ37)。そして、この多重化信号に基づいて送信データが作成されるかまたはこの多重化信号が録音メディアに記録される(ステップ38)。

【0056】このようにして生成された多重化信号が再生装置2側で再生される場合には、操作者によって再生速度(1倍速または2倍速)が設定される。入力音声信号のうち、設定された再生速度に対する圧縮率 α が話速制御情報として設定されている区間では、その圧縮率 α で入力音声信号が時間軸圧縮伸長処理された後、再生される。また、入力音声信号のうち、話速制御情報によって削除される区間であると指定されている区間では、入力音声信号が削除される。

【0057】この第3動作例においても図6の第2動作例と同様に、無音区間と判別された場合には(ステップ32でNO)、無音区間の継続長を話速制御情報として生成するようにしてもよい。

【0058】図8は、編集装置1の第4動作例を示している。ここでは、映像と音声とを伴う放送番組の音声信号が入力されているものとする。また、説明の便宜上、再生装置2側で設定される再生速度としては、1倍速再生と2倍速再生との2種があるとする。また、音声信号に多重化される話速制御情報には、1倍速再生用と2倍速再生用とがある。

【0059】まず、音声信号分析部11によって入力音声信号の所定区間毎に入力音声信号のパワー平均値 P が算出される(ステップ51)。次に、パワー平均値 P が所定のしきい値 T_h 以上か否かが判別される(ステップ52)。

【0060】パワー平均値 P が所定のしきい値 T_h 以上($P \geq T_h$)である場合には、当該区間は音声区間であると判別される。そして、フラグ F の状態に基づいて、前回の区間は継続数が所定数 T_{del} 以上の無音区間であったか否かが判別される(ステップ53)。

【0061】前回の区間が、継続数が所定数 T_{del} 以

上の無音区間でない場合($F=0$)には、圧縮率 α が現在設定されている圧縮率に対して単調減少するように設定される(ステップ54)。2倍速再生用の圧縮率 α は、たとえば、 $1/2 \leq \alpha \leq 1$ の範囲内で設定される。圧縮率 α が2倍速再生時の一般的な圧縮率である $1/2$ である場合には、出力音声速度は入力音声速度の2倍となり、圧縮率 α が1である場合には、出力音声速度は入力音声速度の1倍となる。したがって、出力音声速度が入力音声速度の1倍以上で2倍以下となる範囲内で、圧縮率 α が設定される。

【0062】ステップ54で、圧縮率 α が単調減少されるということは、2倍速再生時において、出力音声速度が入力音声速度の1倍以上で2倍以下となる範囲内で、再生時の出力音声速度が速くなるように圧縮率 α が設定されることを意味する。これは、2倍速再生時において、入力音声信号に対する出力音声信号の遅延時間が所定時間以上になるのを防止するためである。

【0063】1倍速再生用の圧縮率 α は、たとえば、 $1 \leq \alpha \leq 3/2$ の範囲内で設定される。圧縮率 α が1である場合には、出力音声速度は入力音声速度の1倍となる。圧縮率 α が $3/2$ である場合には、出力音声速度は入力音声速度の $2/3$ 倍となる。したがって、出力音声速度が入力音声速度の $2/3$ 倍以上で1倍以下となる範囲内で、圧縮率 α が設定される。

【0064】ステップ54で、圧縮率 α が単調減少されるということは、1倍速再生時においては、出力音声速度が入力音声速度の $2/3$ 倍以上で1倍以下となる範囲内で、再生時の出力音声速度が速くなるように圧縮率 α が設定されることを意味する。これは、1倍速再生時において、入力音声信号に対する出力音声信号の遅延時間が所定時間以上になるのを防止するためである。

【0065】次に、1倍速再生速度および2倍速再生速度に応じてそれぞれ設定された圧縮率 α からなる話速制御情報が生成され、コード化部12によってコード化される(ステップ55)。次に、多重化部13によって、当該区間の入力音声信号と話速制御情報とが多重化される(ステップ56)。そして、この多重化信号に基づいて送信データが作成されるかまたはこの多重化信号が録音メディアに記録される(ステップ57)。

【0066】上記ステップ53において、前回の区間が、継続数が所定数 T_{del} 以上の無音区間である($F=1$)と判別された場合には、フラグ F がリセット($F=0$)にされる(ステップ58)。そして、圧縮率 α が現在設定されている圧縮率に対して大きくなるように設定される(ステップ59)。つまり、2倍速再生用の圧縮率 α が $1/2 \leq \alpha \leq 1$ の範囲内で設定される場合には、2倍速再生時において、出力音声速度が入力音声速度の1倍以上で2倍以下となる範囲内で、再生時の出力音声速度が遅くなるように2倍速再生用の圧縮率 α が設定される。また、1倍速再生用の圧縮率 α が $1 \leq \alpha \leq 3$

／2の範囲内で設定される場合には、1倍速再生時に於いて、出力音声速度が入力音声速度の2／3倍以上で1倍以下となる範囲内で、再生時の出力音声速度が遅くなるように1倍速再生用の圧縮率 α が設定される。

【0067】そして、1倍速再生速度および2倍速再生速度に応じてそれぞれ設定された圧縮設定 α からなる話速制御情報が生成され、コード化部12によってコード化される(ステップ55)。次に、多重化部13によって、当該区間の入力音声信号と話速制御情報とが多重化される(ステップ56)。そして、この多重化信号に基づいて送信データが作成されるかまたはこの多重化信号が録音メディアに記録される(ステップ57)。

【0068】上記ステップ52において、パワー平均値Pが所定のしきい値Thより小さいときには($P < T_h$)、当該区間は無音区間であると判別され、無音区間の継続数が算出される(ステップ60)。そして、無音区間の継続数が所定数Tdel以上であるか否かが判別される(ステップ61)。無音区間の継続数が所定数Tdelより少ないときには、フラグFがセット($F = 1$)される(ステップ62)。また、音声信号の時間長さを伸長する区間と判定される(ステップ63)。

【0069】次に、上記ステップ54と同様に、1倍速再生用および2倍速再生用の圧縮率 α が現在設定されている対応する圧縮率に対して単調減少するように設定される(ステップ64)。そして、1倍速再生速度および2倍速再生速度に応じてそれぞれ設定された圧縮設定 α からなる話速制御情報が生成され、コード化部12によってコード化される(ステップ65)。次に、多重化部13によって、当該区間の入力音声信号と話速制御情報とが多重化される(ステップ56)。そして、この多重化信号に基づいて送信データが作成されるかまたはこの多重化信号が録音メディアに記録される(ステップ57)。

【0070】上記ステップ61において、無音区間の継続数が所定数Tdel以上であると判定されたときには、当該区間は削除すべき区間と判定される(ステップ66)。そして、当該区間の入力音声信号を削除区間とする制御情報が生成され、コード化部12によってコード化される(ステップ67)。

【0071】次に、多重化部13によって、当該区間の入力音声信号と話速制御情報とが多重化される(ステップ56)。そして、この多重化信号に基づいて送信データが作成されるかまたはこの多重化信号が録音メディアに記録される(ステップ57)。

【0072】このようにして生成された多重化信号が再生装置2側で再生される場合には、操作者によって再生速度(1倍速または2倍速)が設定される。入力音声信号のうち、設定された再生速度に対する圧縮率 α が話速制御情報として設定されている区間では、その圧縮率 α で入力音声信号が時間軸圧縮伸長処理された後、再生さ

れる。また、入力音声信号のうち、話速制御情報によって削除される区間であると指定されている区間では、入力音声信号が削除される。

【0073】この第4動作例においても図6の第2動作例と同様に、無音区間と判別された場合には(ステップ52でNO)、無音区間の継続長を話速制御情報として生成するようにしてもよい。

【0074】また、図9に示すように、所定数以上の無音区間に挟まれた音声の1文章の長さを測定し、1文章の長さに応じて圧縮率 α を単調減少させるようにしてもよい。つまり、文頭は圧縮率 α が大きく(出力音声速度を遅く)なり、文末にいくに従って圧縮率 α が小さく(出力音声速度を速く)なるように、話速制御情報を作成してもよい。

【0075】また、入力音声のピッチが急激に上昇した地点では圧縮率 α が大きく(出力音声速度を遅く)なり、その後入力音声のピッチが下降するにしたがって、圧縮率 α が小さく(出力音声速度を速く)なるように、話速制御情報を作成してもよい。

【0076】また、入力音声の発声速度を検出し、発声速度に応じて圧縮率 α を決定するようにしてもよい。すなわち、発声速度が速い場合には、圧縮率 α が大きく(出力音声速度を遅く)なり、発声速度が遅い場合には、圧縮率 α が小さく(出力音声速度を速く)なるように、話速制御情報を作成してもよい。発声速度の検出方法としては、単位時間当りの母音の数または母音の継続長を検出する方法、単位時間当りの音声区間と無音区間の割合を検出する方法等がある。

【0077】上記各実施例では、入力音声を音声区間と無音区間とに区別しているが、入力音声を有音区間、無音区間および無音区間に区別するようにしてもよい。これらの判別は、パワー平均値により、まず、有音区間と、無音区間および無音区間とを区別する。次に、無音区間と無音区間との判別は、音声信号の零交差数(零レベルとの単位時間当りの交差数)を計算し、零交差数が所定数以上の区間を無音区間と判別し、零交差数が所定数未満の区間を無音区間と判別する。有音区間に対しては音声信号の時間長さを伸長する区間とし、所定長以上の無音区間を削除区間とし、無音区間は処理しない区間として、話速制御情報を作成する。また、有音区間の場合には、自己相関法を用いて音声のピッチを抽出し、このピッチ抽出情報を話速制御情報として作成してもよい。

【0078】また、音声信号が音楽である場合には、再生装置側で圧縮処理が行なわれないようにするために、話速制御情報として音楽であることを示す情報を入れるようにしてもよい。音楽であるか否かを判別する方法としては、入力音声信号を周波数分析し、高周波数域(4KHz以上)の信号成分と、低周波域(4KHz未満)の信号成分との割合を算出し、高周波数域成分が大きい

場合は音楽として判断する方法がある。

【0079】また、技術情報番組等においては、音声信号を分析することによって所定の技術用語のフレーズを検出し、検出した区間に所定のキーワードを話速制御情報として付加しておき、再生時において操作者が上記キーワードを再生装置に設定した場合には、キーワードに対応する区間（上記技術用語のフレーズ）をゆっくり再生するようにしてもよい。

【0080】なお、上記実施例では、編集装置において音声信号と話速制御情報とが多重化されて出力されているが、音声信号と話速制御情報とを切り替え伝送方式によって出力するようにしてもよい。

【0081】また、再生装置側において、1倍速再生、2倍速再生の他、3倍速再生等が可能な場合にも、圧縮率 α を全ての再生倍率速度ごとに設定して、話速制御情報とし音声信号に付加することが好ましい。

【0082】

【発明の効果】この発明によれば、予め送信データ、録音メディア等に話速を制御するための情報を入れておき、受信側または再生側において、制御情報に基づいて話速を制御できる話速変換システムが得られる。

【図面の簡単な説明】

【図1】話速制御システムの構成を示すブロック図である。

【図2】編集装置の構成を示すブロック図である。

【図3】再生装置の構成を示すブロック図である。

【図4】再生装置の話速制御部の構成を示すブロック図である。

【図5】編集装置の第1動作例を示すフローチャートである。

【図6】編集装置の第2動作例を示すフローチャートである。

【図7】編集装置の第3動作例を示すフローチャートである。

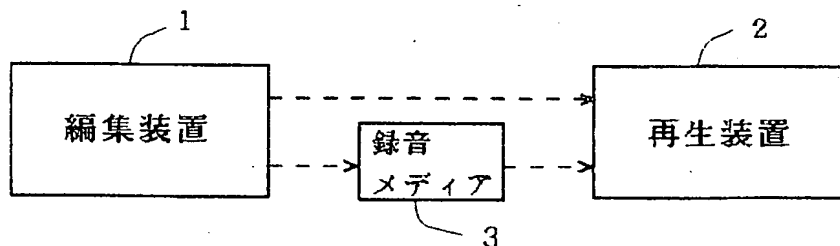
【図8】編集装置の第4動作例を示すフローチャートである。

【図9】所定長以上の無音区間に挟まれた1文章の区間を示す模式図である。

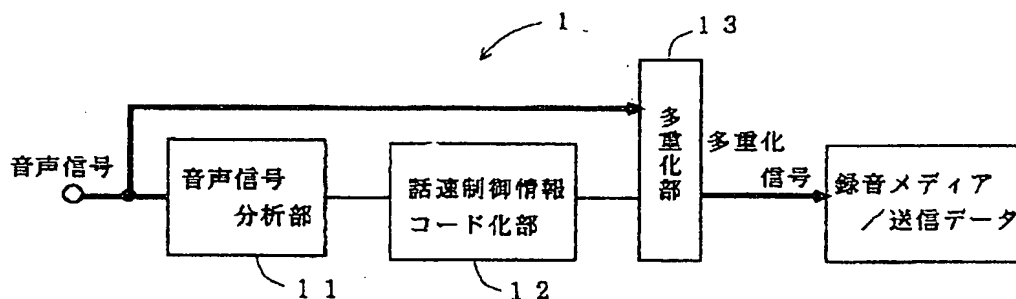
【符号の説明】

- 1 編集装置
- 2 再生装置
- 11 音声信号分析部
- 12 話速制御情報コード化部
- 13 多重化部
- 21 復調部
- 22 話速制御部
- 31 話速制御情報解析部
- 32 制御情報同期化部
- 33 話速変換部

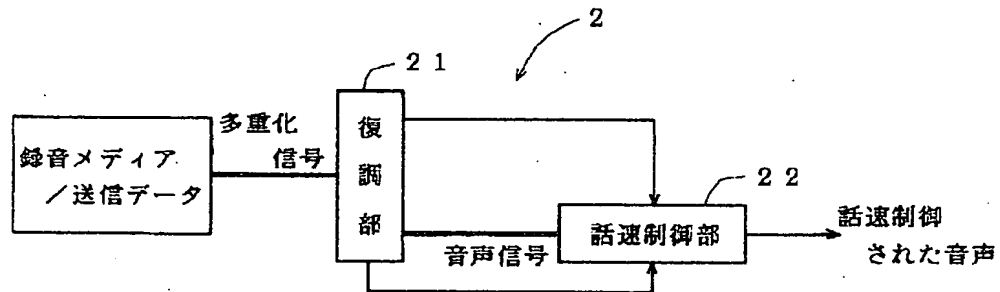
【図1】



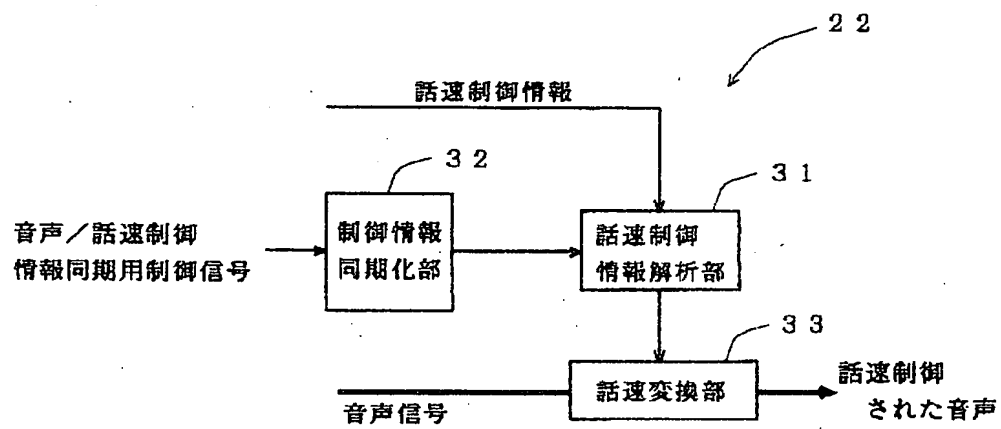
【図2】



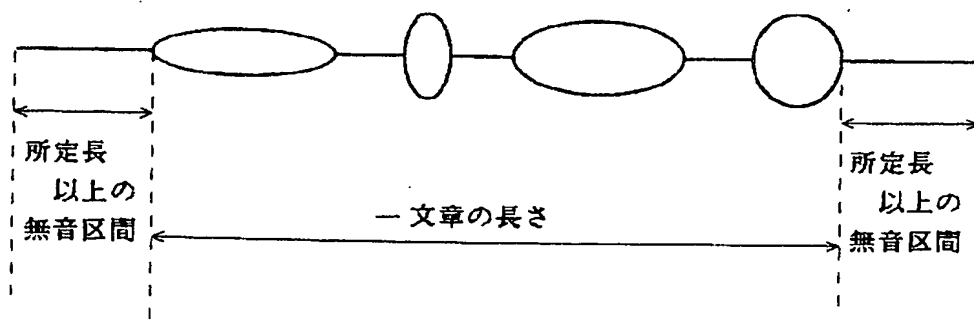
【図 3】



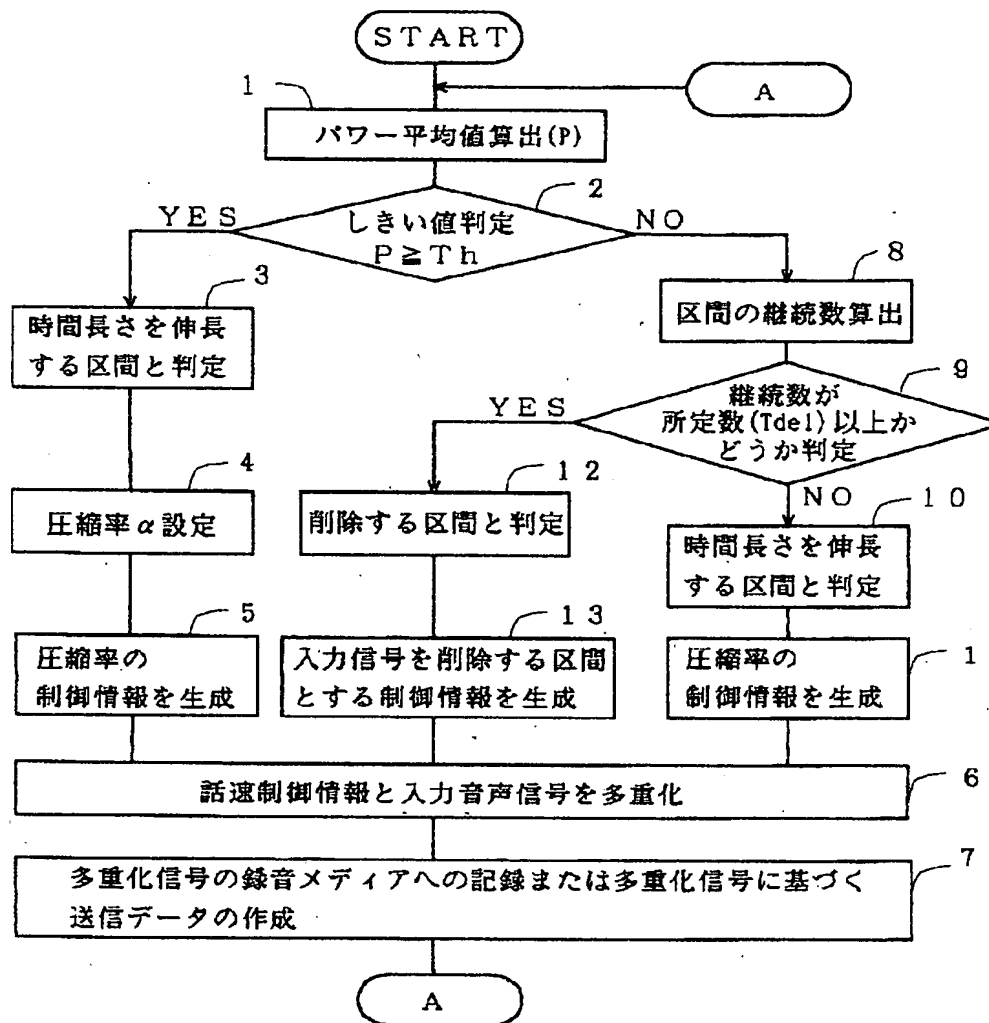
【図 4】



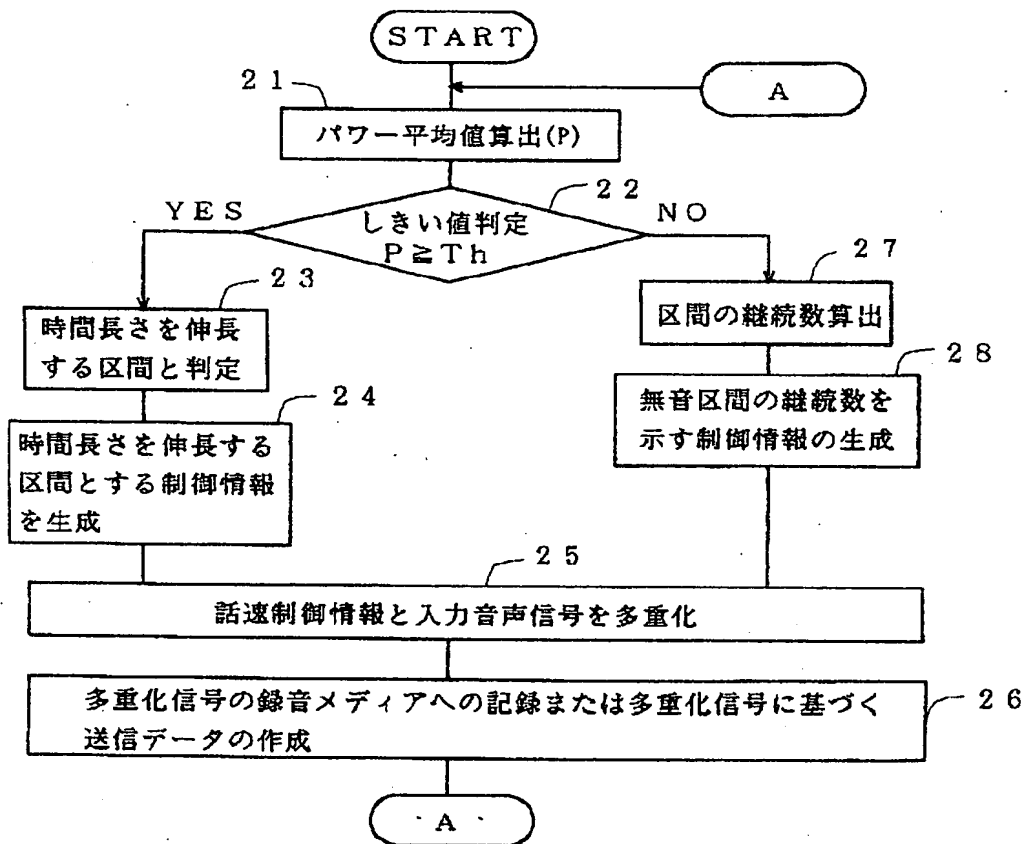
【図 9】



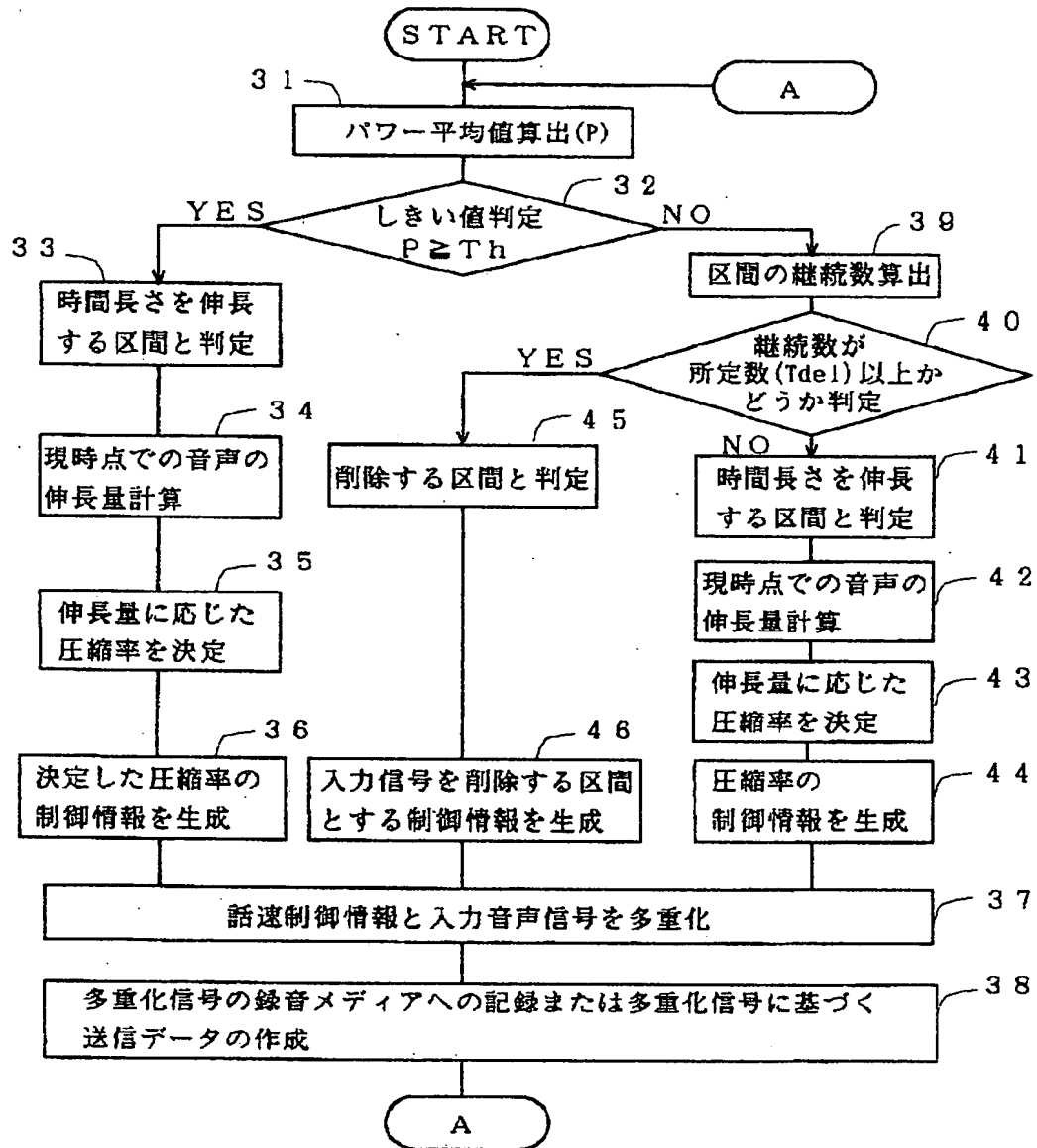
【図5】



【図 6】



【図7】



【図8】

